AWARD NUMBER: W81XWH-14-1-0473

TITLE: Genomic Diversity and the Microenvironment as Drivers of Progression in DCIS

PRINCIPAL INVESTIGATOR: EUN-SIL HWANG/CARLO MALEY

CONTRACTING ORGANIZATION: DUKE UNIVERSITY
Durham, NC 27710

REPORT DATE: October 2016

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

1. REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2016	Annual	30Sep2015 - 29Sep2016

4. TITLE AND SUBTITLE

Genomic Diversity and the Microenvironment as Drivers of Progression in DCIS

5a. CONTRACT NUMBER
W81XWH-14-1-0473

5b. GRANT NUMBER
BC132057

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

E.S. Shelley Hwang

E-Mail: shelley.hwang@dm.duke.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Duke University
2200 W Main St Ste 710
Durham NC 27708-4677

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT: The project is designed to test whether genetic and/or tumor environmental heterogeneity is a driving force in progression of breast DCIS. Our project, a collaboration between Duke and ASU, has made substantial progress on all 4 aims and we met our 24 month milestones. Primary achievements for 24 months are: 1) Continued Case and control identification through extensive database and medical record searching at Duke, 2) Development of methods for isolating DNA from archival DCIS lesions, 3) Deep and comprehensive full exome sequencing for 20 cases from 20-160ng of DNA isolated from these archival FFPE specimens, 4) Comparison of analytic methods to characterize somatic mutations from this full exome sequencing, 5) Application of sequencing library DNA to Illumina SNP arrays for copy number assessment 6) Development of dual immune-staining on DCIS lesions using 6 pairs of antibodies, 7) Sharing of images from these stains with collaborators for quantitative analysis, 8) Identification of a series of upstaged DCIS cases for the radiology aim, 9) Development of image analysis methods for digital mammograms, 10) Approval of both the TBCRC and Duke IRB protocol for the validation aim to initiate collection of DCIS that either did or did not progress to invasive cancer, 11) Full integration of team members over the past year via frequent conferencing, face to face meetings, and constant communication. This multi-disciplinary progress puts our group into an ideal position to fully implement the aims of the project and reach our year 3 and 4 goals.

15. SUBJECT TERMS
DCIS, intra-tumor heterogeneity, genetic diversity, phenotypic diversity, somatic evolution, microenvironment, mammographic biomarkers

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	15	

Table of Contents

	Page
Introduction……..	**1**
Body……..	**1**
Keywords.....……..	**1**
Impact.....……	**11**
Changes/Problems……	**11**
Publications……	**12**
Participants & Other Collaborating Organizatons……	**12**

1. INTRODUCTION

Ductal carcinoma in situ (DCIS) of the breast is an increasingly common diagnosis that is related to aggressive screening patterns (mammography). This "pre-invasive" lesion may progress to invasive cancer, but does so at a relatively low frequency. Nonetheless, it is commonly treated with extensive surgery, radiation, and hormonal therapy even though most of these lesions would never progress to invasive cancer. Thus, there is a pressing clinical need to stratify the risk of DCIS tumors into those in need of intervention and those that can be safely monitored without intervention. Our project is designed to address this need by characterizing the evolvability of DCIS, detecting those that have a high likelihood of evolving to malignancy versus those that are likely to remain indolent.

2. KEYWORDS

DCIS, cancer progression, intra-tumor heterogeneity, genetic diversity, phenotypic diversity, somatic evolution, microenvironment, mammographic biomarkers

3. ACCOMPLISHMENTS

What were the major goals of the project?

Aim 1. Determine whether genetic diversity of DCIS is greater in DCIS with adjacent invasive disease compared to DCIS without progression. Diversity measures must be derived from geographically distinct areas of tumor. Genetic divergence of the DCIS component of tumors will be measured based on exome sequencing and SNP arrays run on two separate regions of the tumor, as well as normal tissue, in patients with DCIS either with or without invasion to determine the association between genetic diversity and progression to malignancy. Genetic diversity will be measured by the genetic divergence between the tumor samples, that is, the proportion of the genome that differs between the two samples from the same tumor.

24 Month Milestones:

Protocol preparation, IRB submission and approval: **Completed** (Duke eIRB Pro00054515, initial Duke approval, 5/27/2014 and renewed for the current year), DOD IRB approval in place.

Case identification and tissue block selection: **Ongoing; on schedule**. Through a variety of available databases, we identified a large number of cases and controls with tissue available in the Duke Pathology archives. Each potential case and control requires extensive chart and pathology review in order to determine final eligibility and usability. For example, there is sufficient amount of the DCIS lesion (>2mm size) for isolation and DCIS is not too close to invasive cancer (it extends outside the invasive component). There must be two blocks with DCIS present that are >0.8cm apart. To date we have identified 42 cases.

Sectioning of tissue blocks: **Ongoing; on schedule**. New sections from candidate paraffin blocks are cut, stained with H&E, reviewed by the study pathologist. Remaining sections from candidate blocks (containing a sufficient amount of the DCIS lesion of interest) are used for macro-dissection and subsequent DNA extraction. Additional sections (every other one) are also stored for immunohistochemical (IHC) analysis of key measures of tumor and micro-environmental heterogeneity. These slides are scanned for analytic and archival purposes. This process has been fully implemented and we are moving through both cases and controls in this manner.

DNA extraction of test cases: **Completed**.

Exome sequencing of test cases: **Completed.** We have investigated a number of platforms and collaborators for the DNA sequencing and SNP analysis. Since we are working with small amounts of FFPE DNA, standard methodologies do not readily apply. We have settled on the Genome Center at Washington University. Wash U. has developed cutting-edge methods for producing high quality data from these specimens. In addition to full-exome capture, the method employs additional enrichment for a panel of 83 genes to ensure high coverage of the most commonly altered breast cancer driver genes. Over the past several months, Wash U. sequenced 20-160ng from 81 individual DNA samples derived from 26 subjects (germ line sample plus 2 DCIS containing samples) and returned the data to us for analysis. They were able to derive interpretable sequence data from 20-160ng of FFPE DNA with qualities summarized in Figure 1.

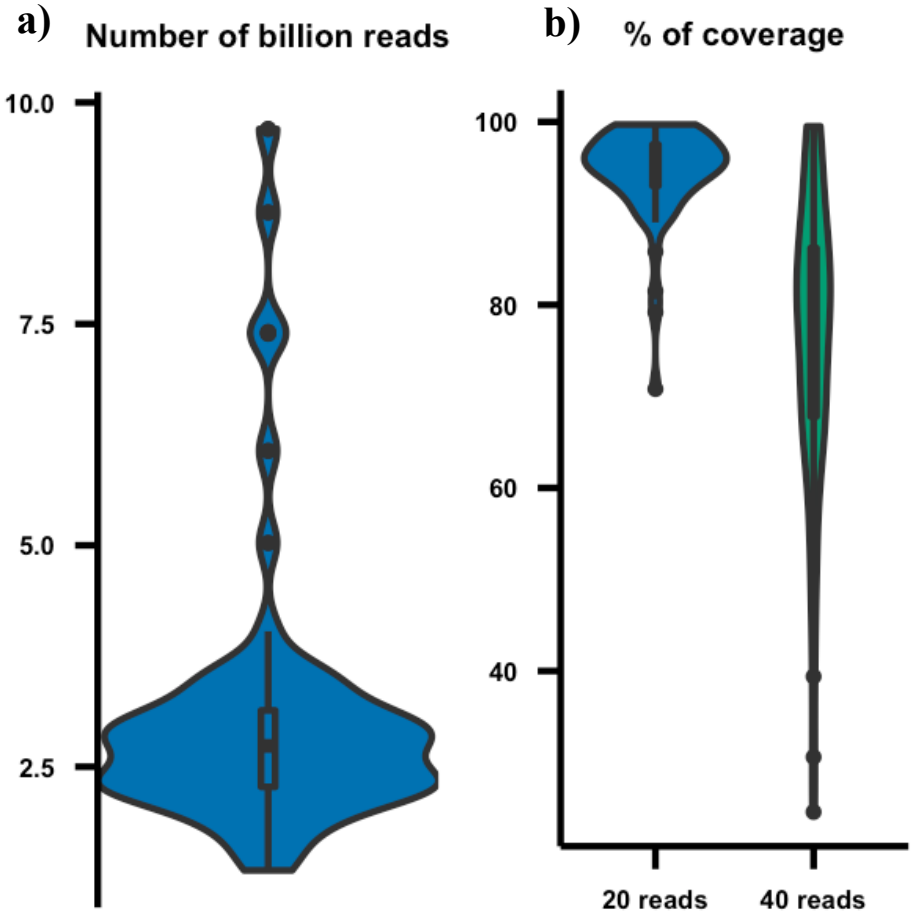

Figure 1. Quality control summary. A) Violin plot showing the distribution of number of reads for all samples used in the analysis. B) Violin plots showing the distribution of the percentage of the exome covered by two different depths (20 and 40 reads).

SNP arrays: In order to better estimate copy number variation (CNV), we are also analyzing DNA from the two areas of DCIS from each case using high density single nucleotide polymorphism (SNP) arrays. We are using the human Omni Express array from Illumina to accomplish this aspect of the project. Since DNA from the primary samples is limiting (macrodissected DCIS), we have been testing whether sequencing libraries generated for exome sequencing can be directly applied to these arrays.

Development of a pipeline for identification of somatic genetic alterations: **Completed.** In order to assess and minimize artefacts induced by the FFPE procedure and the small amounts of DNA obtained from FFPE samples we developed a strategy based on 9 sequencing technical replicates. We sequenced the DNA extracted from different tumor samples twice to obtain a series of 9 independent pairs of technical replicates. Consequently, each sequence should be identical to its replicates. Therefore, any differences between the pair of technical replicates is due to current technical limitation of NGS technological platform mainly induced by limited and degraded DNA templates. We developed a pipeline that automatically explores more than 600K options of variant calling and posterior variant filtering, looking for the one that minimizes the divergence of those technical replicates. We explored different parameters and optimization criteria, and used the best to obtain the final filtering options we applied in the analysis of our pilot cohort. Although the pipeline has been completed and is fully functional, we will refine our sequencing analysis pipeline.

Calculation of genetic diversity scores for the pilot cohort: **Completed**. The main purpose of the research project is to determine the heterogeneity between samples. Nevertheless, false positives directly increase the heterogeneity estimation. In order to reduce the false positive rate, we applied stringent exclusion criteria. However, if we filtered the two samples to compare with the same criteria we would have high chances of missing true variants because the sequences have different level of coverages and quality. We developed a two-step comparison strategy in order to solve this issue, by which we compare one of the samples filtered and the other unfiltered and the other way around. Then, we calculate the union of common variants coming from the filtered side of each comparison. By applying this strategy, we were able to identify the best filtering parameters and their values, allowing us to obtain an average level of similarity of 0.81% ± 0.12 SD among the technical replicates.

Using this approach, and the best filtering strategy resulting from our technical-replicate based algorithm, we analyzed the pilot cohort. The results are summarized in the plots below (Figure 2). We did not find significant differences in the mean divergence (Figure 2a) or number of mutations (Figure 2b) between the two cancer types. Nevertheless, we did find strong differences in terms of divergence variance between groups, which may indicate that they are driven by different evolutionary forces. Additionally, we found differences close to be significant in terms of fold differences in number of mutations between regions (Figure 2c). Moreover, the mean number of somatic non-synonymous mutation is much higher in the DCIS with adjacent invasive (Figure 2d), which may indicate a difference in selective pressures between cancer types. This difference also approaches significance, which is encouraging given the small amount of cases in the pilot study.

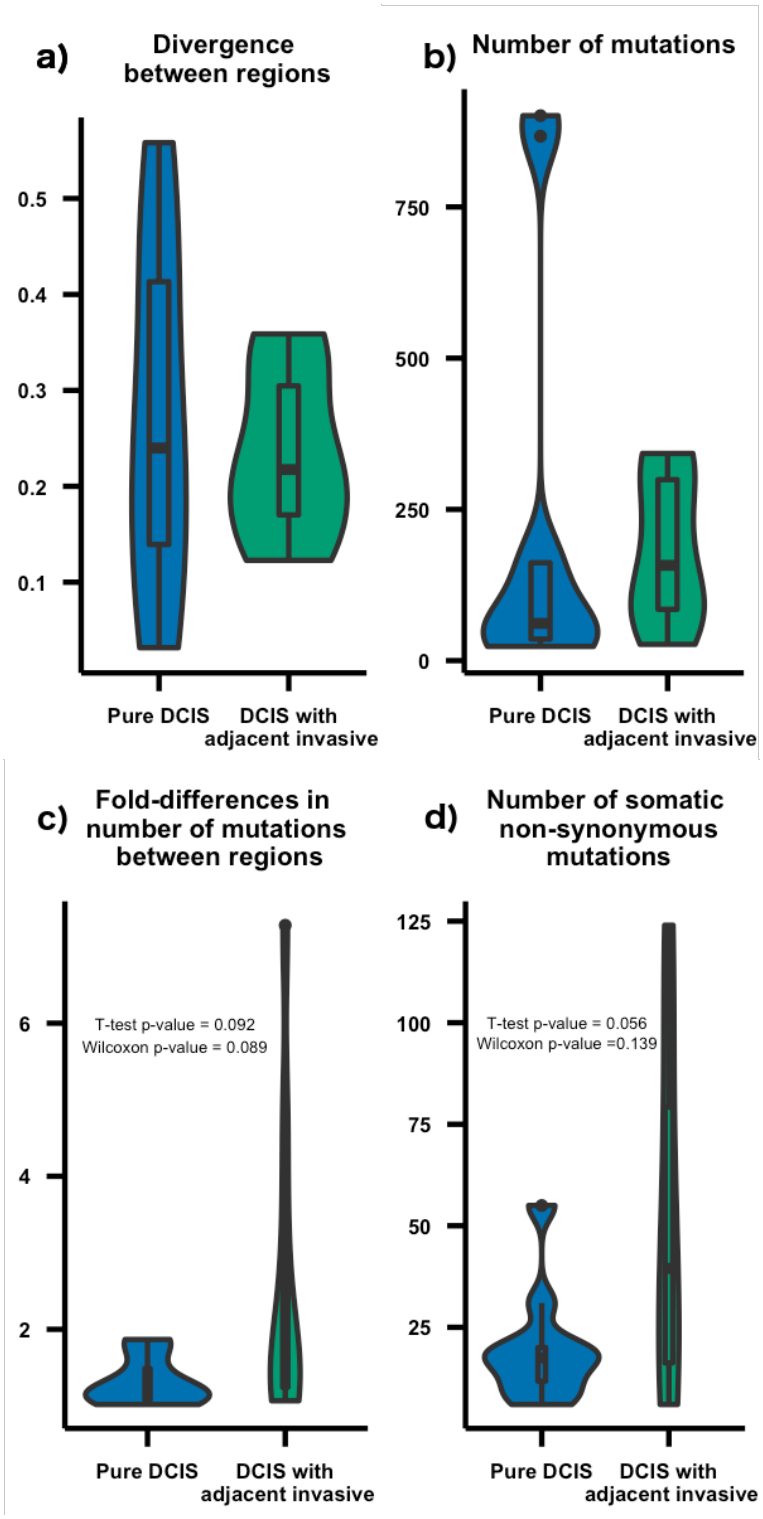

Figure 2. Summary of the results of the pilot cohort. Violin plots showing the distribution for the two different cancer types of divergence, number of mutations, fold-differences and number of somatic mutations in a), b), c) and d), respectively.

Aim 2. Determine whether phenotypic diversity of DCIS and the tumor microenvironment (TME) is greater in DCIS with adjacent IDC compared to DCIS without IDC. Since genomics is not the sole driver of tumor behavior, we will phenotypically characterize DCIS and its microenvironment including markers of hypoxia, migration, proliferation, matrix organization, and immune signaling in the same samples used in Aim 1. We will employ automated image analysis to compute microenvironmental divergence to determine if specific components of the TME, or the divergence between TMEs from the same tumor, differs between DCIS with and DCIS without adjacent IDC.

In the past 10 months, we have finalized a series of analytes and parameters that are intended to capture a series of important elements of the phenotypic diversity of DCIS. These elements include the presence and distribution of cell types (malignant epithelia, lymphocytes, and stroma) and expression of proteins that are associated with oncogenic and environmental processes. To evaluate these elements, we are using a combination of expert scoring and automated image analysis. Further, expert scoring is being used to guide, train, and evaluate the image analysis so these analyses will be cross-informative.

The first set of 21 cases (10 pure DCIS and 11 DCIS concurrent with invasive cancer) has been evaluated using both expert scoring and image analysis for cell content. This set (two areas or blocks from each case) has also been stained for the set of phenotypic expression markers (Table 1) and scored by expert analysis. All results are recorded and tabulated in a shared study database.

Table 1: Phenotypic Markers of Heterogeneity & Pathology Scoring

Double Stains	Functional Category	Cell Type	Scoring	
ALDH1A1	Stem Cell Marker	Epithelia	Intensity + Dist	
Ki-67	Proliferation	Epithelia	Distribution	
COL15A1	Basement Membrane	BM	Presence around DCIS	↓p=0.027
ESR1	Hormone Signaling	Epithelia	Int + Dist	
Phospho-FAK	Cell Adhesion	Epithelia	Int + Dist	
CD68	Macrophage	Macrophage	Distribution	
CA9	Hypoxia	Epithelia	Int + Dist	↓p=0.045
FOXP3	T Regulatory Cells	Lymphocyte	Distribution	
ERBB2	Oncogenic Signaling	Epithelia	Int + Dist	
P63	Basal Cells	Myoepith	Presence around DCIS	
RANK	Inflammatory Signaling	Epithelia	Int + Dist	↓p=0.046
PGR	Hormone Signaling	Epithelia	Int + Dist	
Single Stains				
GLUT1	Glucose Transport	Epithelia	Int + Dist	↓p=0.059
CD31	Blood Vessels	Endothelia	Distribution	
Rho A	Motility	Epithelia	Int + Dist	

These 15 markers represent a range of processes, cell types, structures, and environmental influences on the tumor. Based on their biologic significance, we have developed methods for scoring each of the markers summarized in the table above. In the next funding cycle, we will continue to stain and analyze all cases for the study in this manner. We will continue to refine image analysis,

particularly as it relates to quantitative scoring of the immunohistochemical staining. We will monitor and evaluate concordance between expert and image analysis.

We have brought a new collaborator into the team, Dr. Yinyin Yuan from the Center for Evolution and Cancer at the Institute for Cancer Research in London. Dr. Yuan is an expert in computational image analysis of histological sections of breast cancer, and the application of ecological and other spatial statistics to those images [1-4]. She and her group will provide quantitative analyses of immunohistochemical stained sections to evaluate tumor heterogeneity.

24 Month Milestones:
IHC staining of candidate markers (test cases): We have obtained and characterized a series of antibodies representing our initial targets including ER, PR, KI-67, COL15A1, RHOA, RAC, CA9, HIF1a, FOXP3, and cleaved Caspase 3. We have piloted dual staining for sets of these antibodies on test cases of breast cancer and will soon be staining for these antigens on DCIS cases and controls. Dual staining conditions will be optimized in collaboration with Dr. Yinyin Yuan's lab who will perform the automated, quantitative scoring and analysis of the stained tissues.

Scan IHC results for Automated image analysis (AIA)

Automated image analysis (AIA) of tumor and stromal markers of heterogeneity: Dr. Yuan's team is adapting their algorithms for dual staining. They already have successfully analyzed both clustering of cell types [2, 3], and co-localization (interleaving) among different cell types (manuscript under review).

We have performed IHC staining of the pilot cohort with the following markers: we have used single stains for CD31, GLUT1, RhoA; we have used double stains for the following pairs KI-67 and ADLH1A1, Phospho-FAK and CD68, COL15A1 and ESR1, CA9 and FOXP3, ERBB2 and P63, RANK and PGR. Dual staining conditions have been optimized in collaboration with Dr. Yinyin Yuan's lab to enable automated, quantitative scoring and analysis of the stained tissues.

We have performed manual scoring of the IHC markers in the pilot cohort. The expert pathology review revealed that DCIS with adjacent IDC exhibits less hypoxia, as revealed by CA9 scoring, lower GLUT1 expression and less inflammatory signaling (Table 1).

We have performed Automated image analysis (AIA) on the H&E slides in order to automatically score epithelial cells, lymphocytes and fibroblasts. The DCIS microenvironment has been further analyzed to detect epithelial and lymphocytic hotspots and to quantify the microenvironment heterogeneity and co-localization of different cell types (Figure 3a, 3b, H&E).

We have developed AIA methods for automated scoring of dual-stained IHC and for detection of epithelial cells, lymphocytes and fibroblasts in IHC slides. We have demonstrated the scoring on pilot cohort for Ki67/ADLH1 dual-stains (Figure 4a, 4b, IHC).

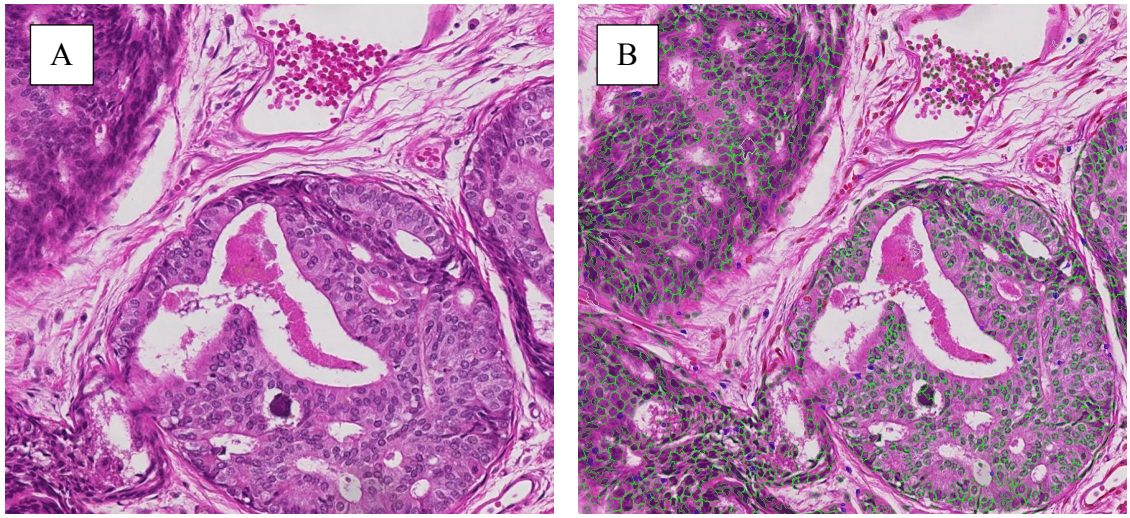

Figure 3: A and B H&E slide

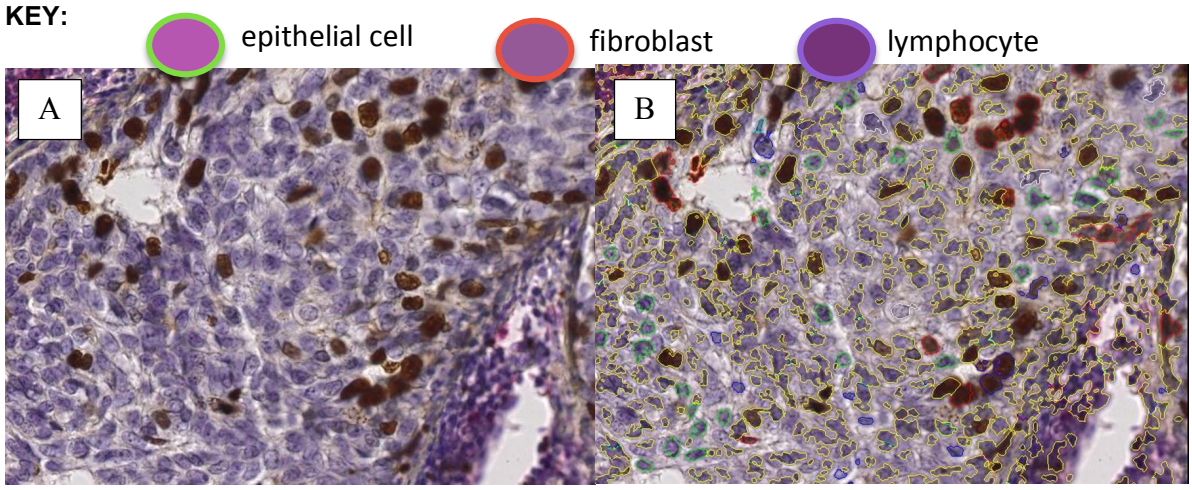

Figure 4: A and B, Dual Stained IHC for Ki67 & ALDH1 (proliferation & stemness)

Aim 3. Create and test a computational learning algorithm to compare mammographic characteristics and diversity measures in pure DCIS compared to DCIS with IDC. A weighted computational algorithm using mammographic features of lesional and stromal characteristics as well as heterogeneity measures derived from Aims 1 and 2 will be constructed. The tool will be designed to allow for radiologic discrimination between good and poor prognosis DCIS, and will be evaluated in a validation set.

24 Month Milestones:
We re-implemented the computer vision algorithm to be 10 times faster and more robust. Based upon a greatly expanded set of 105 images from 55 preliminary training cases, the new algorithm is comprised of 3 main steps: (1) mammogram signal enhancement using contrast-limited adaptive histogram equalization, morphological operations, and top-hat filtering, (2) coarse microcalcification detection using locally adaptive thresholding, and (3) false positive reduction using morphological rules and weighted graph clustering.

Independent of the above preliminary segmentation training cases, we have now completed the identification of the DCIS cross-validation and test data sets. We queried our EMR system to identify 5,300 initial candidates of DCIS cases undergoing needle core biopsy, which were then semi-automatically filtered using the inclusion/exclusion criteria to yield 198 subjects. We selected half randomly for the cross-validation set, comprised of 74 cases of pure DCIS vs. 25 cases that were upstaged to invasive disease at definitive surgery. The other half of 99 testing subjects have been set aside for aim 3b work.

For the first group of 99 cross-validation DCIS cases, we have automatically extracted 3 types of computer vision features: (1) shape features to describe heterogeneous morphology and size of calcifications and clusters, (2) topology features based on relations between calcifications from weighted graphs, and (3) texture features based on calcifications vs. background pixel values and statistical measures of Gray Level Co-occurrence Matrices. There are 13 cluster-level features and 100 individual calcification features or a total of 113 features per lesion.

We have begun the process of statistical analysis slightly ahead of schedule (aim 3b, originally scheduled for months 30-36). We have collected clinical findings from pathology reports, as well as a radiologist observer study to describe mammographic appearance of the DCIS lesion and provide subjective assessment of the likelihood of upstaging, as shown in Table XXX. A second radiologist will complete this study before the end of the current 24 month period. We have also begun the statistical analysis of the 113 computer vision features. Figure XXX demonstrates two exemplar computer vision features that are selected frequently in leave-one-out cross-validation sampling.

Table 2: Comparison of histologic and mammographic features between DCIS and invasive groups

Feature		DCIS n=74	Invasion n=25	p-value
Histology	**Nuclear grade (1-3)**	2.51	2.58	p = 0.6044
	1	6	3	
	2	27	6	
	3	41	16	
	Subtype of DCIS			p = 0.5250
	Comedo	36	14	
	Non-Comedo	38	11	
BI-RADS Mammography	**Age**	59.8	58.2	p = 0.5330
	Size of lesion			
	Area (mm^2)	210.3	369.0	p = 0.2257
	Axis (mm)	16.7	24.8	**p = 0.0496 ***
	Morphology of calcifications			p = 0.0704
	Low Risk(Typically benign)	0	1	
	Medium Risk (Amorphous/Coarse heterogeneous)	41	7	
	High Risk (Fine pleomorphic/Fine linear)	33	17	

	Distribution of calcifications			p=0.6653
	Regional	2	0	
	Segmental	5	3	
	Linear	2	1	
	Clustered	65	21	
	BI-RADS level of suspicion			**p = 0.0247 ***
	4a	40	7	
	4b	17	8	
	4c	14	8	
	5	3	2	
	Radiologist's score of being invasive	14.5	21.0	**p = 0.0052 ***
**Difference for the comparison was statistically significant.*				

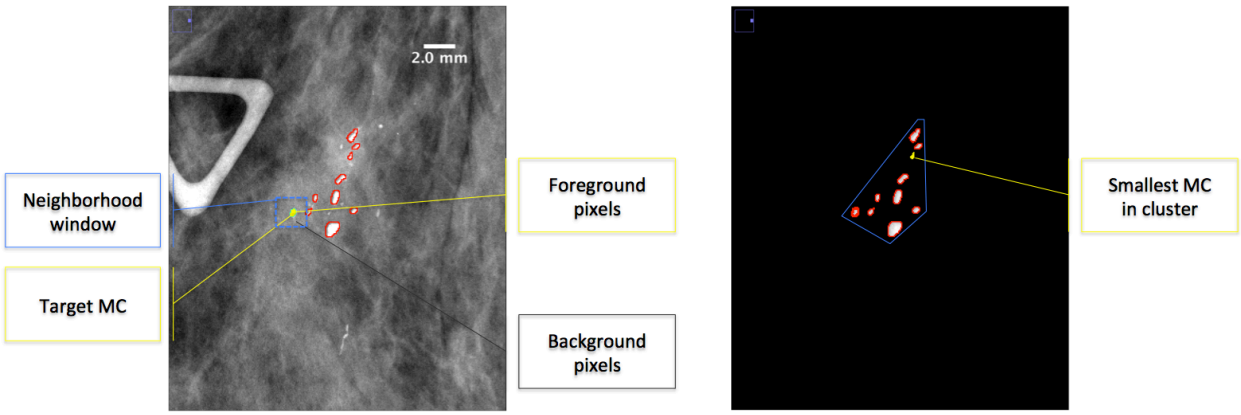

Figure 5. Two exemplar computer vision features: (a) standard deviation of background pixel intensities around microcalcifications, (b) minimum area of individual microcalcifications in a cluster.

Aim 4. Test the predictive performance of the best diversity measures in an independent validation set of pure DCIS with and without subsequent invasive recurrence. Genotypic and phenotypic measures of diversity derived from Aims 1-2 will be applied to an independent case-control, longitudinal, tissue bank of DCIS with and without invasive recurrence to validate their utility. Cases will be obtained through the Translational Breast Cancer Research Consortium (TBCRC), a breast cancer research collaborative of 18 NCI-designated Comprehensive Caner Centers.

The TBCRC protocol has been approved, contracts are being drawn up with the participating sites, the REDCap online data entry forms are being finalized, and accrual will begin on this validation before the beginning of the next budget period. In the next year of funding, we will accrue cases of

pure DCIS that are long term disease free or recurred with metastatic cancer. Slides will be shipped to Duke for macrodissection for DNA analysis and for immunodetection of phenotypic heterogeneity.

What was accomplished under these goals?

Our primary goals have been met including, most importantly, identifying the most efficient method of sequence generation from small amounts of fixed DNA. We have acquired the radiology imaging data sets and established the computer vision algorithms for their analysis. Further, based on our databases, we are confident of accruing sufficient cases and controls at Duke to fulfill the goals of the project. Overall, we are in excellent position to complete the proposed work in the project period along the time line that was provided.

What opportunities for training and professional development has the project provided?

We hired several new post-doctoral fellows in the previous year to continue expanding our analysis. Bibo Shi has been acquiring new skills in medical image analysis and learning about the complexities of breast cancer diagnosis.

How were the results disseminated to communities of interest?

We reported the early sequencing results at the San Antonio Breast Cancer Symposium in December 2015. We have two DCIS abstracts based on aims 1 and 2 submitted to the San Antonio Breast Cancer Symposium in December 2016.

We have also submitted two abstracts based on the Aim 3 results to the SPIE Medical Imaging Conference to be held in February 2017. If accepted, those will each be published in the form of a full-length conference proceedings paper. Those results have been combined into a paper that will be submitted to a peer-review journal by the end of the 24 month period.

What do you plan to do during the next reporting period to accomplish the goals?

Aim 1: We will continue to identify potential cases and controls through Duke Pathology archives and databases and screen for eligibility. Diagnostic slides from candidate subjects will be evaluated by our study pathologist to determine if there is sufficient material to work with and ones that pass this metric will be included in the study. New unstained slides will be ordered from these cases for macrodissection and immunohistochemical staining. DNA extracted from these slides will be exome sequenced and applied to SNP arrays. Returned data from these assays will be analyzed using our current pipeline in order to scale up from the pilot study to a study with a larger sample size. Moreover, we will investigate the biological differences between the most common variants of the two different tumor types. We will also continue to improve our sequencing analysis pipeline by analyzing additional technical replicates. We will report this novel method of analysis of genomic sequence from small amounts of DNA extracted from FFPE samples in a methods manuscript.

Aim 2: We have analyzed cases and controls using a series of antibody stains described in the proposal. Scanned images of these stained slides will be provided to Dr. Yuan for image analysis and quantification. Dr. Yuan's team will adapt their algorithms to quantify dual stained slides. Heterogeneity of expression of these protein markers associated with the tumor, basement membrane, vasculature, and immune infiltrate will be incorporated into measures of genetic heterogeneity. This will be performed on an additional 80 cases over the next budget period.

Aim 3: Using the 99 cross-validation cases, we will focus on extracting computer vision features that specifically pertain to heterogeneity across the lesion and the image. This will complete Aim 3a. We will continue work on Aim 3b to develop imaging-only predictive models using the proposed machine learning techniques.

Aim 4: This multicenter validation arm of the project is set up through the Translational Breast Cancer Research Consortium (TBCRC), a collaborative group set up to conduct innovative and high-impact breast cancer clinical trials. Eleven (12 including Duke) external sites have been enlisted and they are currently obtaining IRB approval at their local IRB. The contracts with each site have been drawn up and execution is based on site IRB approval. Each site will supply approximately 10-12 cases and controls including unstained sections from two DCIS blocks and one germ line block; pathology and imaging data to validate results from Aims 1-3.

The following centers have agreed to participate in this study:

Baylor College of Medicine
Dana Farber/Harvard
Duke University
Georgetown
Indiana University
Mayo Clinic
MD Anderson
Montefiore
University of Chicago
University of North Carolina
University of Pittsburgh
University of Washington/Fred Hutchinson Cancer Center

4. IMPACT

Successful completion of this project will lead to a variety of biomarkers (genetic, IHC and radiographic) to distinguish high risk from low risk DCIS. This would reduce patient suffering and conserve clinical resources for the women with low risk DCIS, and focus management efforts and clinical resources on women with high risk disease, potentially justifying the risks of interventions. As the project is in its initial stages, these important impacts await in the future.

5. CHANGES/PROBLEMS

Changes in approach and reasons for change

There have been no changes in approach.

Actual or anticipated problems or delays and actions or plans to resolve them

So far the problems that have emerged have been primarily technical. Full exome sequencing from small amounts of FFPE tissue is at the limit of current technical practice. Further, analyzing these data is also a challenge. We are now confident in our ability to generate high coverage and high depth sequencing data from as little as 20ng of FFPE DNA. We are also performing technical replicates to determine the reproducibility and noise that is in the system. These data are now being analyzed and will guide the eventual analytic paradigm that will be used going forward.

In order to evaluate heterogeneity within a tumor, we require that there are as few normal cells as possible in the extraction. We initially evaluated laser capture microdissection and found that the DNA yields were insufficient to acquire comprehensive sequence data. Therefore, in conjunction with our study pathologist, we are now routinely marking slides for **macro**dissection which provides sufficient DNA and excellent purity. We are currently developing our automated imaging analyses of dual stained tissue sections with Dr. Yuan and her lab.

6. PRODUCTS

Publications

1. Walther, V., Hiley, C.T., Shibata, D., Swanton, C., Turner, P.E., and **Maley, C.C.**: Can oncology recapitulate paleontology? Lessons from species extinctions. Nature Reviews Clinical Oncology, 12:273-285, 2015. doi:10.1038/nrclinonc.2015.12 Published. Acknowledged federal support.
2. Caulin, A.F., **Maley, C.C.:** Solutions to Peto's Paradox Revealed by Mathematical Modeling and Cross-Species Cancer Gene Analysis. Philosophical Transactions of the Royal Society of London B, 370 (1673):20140222. Published. Acknowledged federal support.
3. Aktipis, C.A., Boddy, A.M., Jansen, G., Hibner, U., Hochberg, M.E., **Maley, C.C.**, Wilkinson, G.S.: Cancer across the tree of life: Cooperation and cheating in multicellularity. Philosophical Transactions of the Royal Society of London B, 370 (1673):20140219. Published. Acknowledged federal support.
4. Noemi Andor, Trevor A. Graham, Marnix Jansen, Li C. Xia, C. Athena Aktipis, Claudia Petritsch, Hanlee P. Ji, Carlo C. Maley: Pan-cancer analysis of the extent and consequences of intra-tumor heterogeneity. Under review at Nature Medicine. Acknowledged federal support.
5. Carlo C. Maley, Konrad Koelble, Rachael Natrajan, Athena Aktipis and Yinyin Yuan: An ecological measure of immune-cancer colocalization as a prognostic factor for breast cancer. Under review at Breast Cancer Research. Acknowledged federal support.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?

Co-PI: Dr. Shelley Hwang (M.D., M.P.H.): Duke University (no change)
Co-PI: Dr. Carlo C. Maley (PhD.): Arizona State University (no change)

Co-Investigators:
Dr. Jeffrey Marks (PhD.): Duke University (no change)
Dr. Lorraine King (PhD): Duke University
Dr. Joseph Geradts (M.D.): Duke University (departed during year one)
Dr. Allison Hall (M.D.): Duke University, replacing Dr. Geradts.
Dr. Joseph Lo (Ph.D.): Duke University (no change)
Dr. Jay Baker (M.D.): Duke University (no change)
Dr. Yin Yin Yuan (PhD): Institute for Cancer Research, UK
Dr. Lars Grimm (M.D.): Duke University.
Dr. Trevor Graham (Ph.D.): Barts Cancer Institute, Queen Mary University of London (no change)
Dr. C. Athena Aktipis (Ph.D.): Arizona State University (no change)
Dr. Shane Jensen (PhD.): University of Pennsylvania (departed during year one)

Post-Docs:
Dr. Mengyu Wang (PhD): Duke University (departed during year one)

Dr. Violet Kovacheva (PhD): Institute for Cancer Research, UK
Dr. Bibo Shi (PhD): Duke University
Dr. Angelo Fortunato (PhD): Arizona State University
Dr. Diego Mallo (PhD): Arizona State University